



## A Randomized Controlled Trial Evaluation of an After-School Pro-Social Behavior Program in an Area of Socioeconomic Disadvantage

O'Hare, L., Biggart, A., Kerr, K., & Connolly, P. (2015). A Randomized Controlled Trial Evaluation of an After-School Pro-Social Behavior Program in an Area of Socioeconomic Disadvantage. *The Elementary School Journal*, 116(1), 1-29. DOI: 10.1086/683102

### Published in:

The Elementary School Journal

### Document Version:

Publisher's PDF, also known as Version of record

### Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

### Publisher rights

Copyright 2015 by the University of Chicago. All rights reserved.

### General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

---

# A RANDOMIZED CONTROLLED TRIAL EVALUATION OF AN AFTER-SCHOOL PROSOCIAL BEHAVIOR PROGRAM IN AN AREA OF SOCIOECONOMIC DISADVANTAGE

---

## ABSTRACT

A randomized controlled trial was used to evaluate the effects of a prosocial behavior after-school program called Mate-Tricks for 9- and 10-year-old children and their parents living in an area of significant socioeconomic disadvantage. The children were randomly assigned to an intervention ( $n = 220$ ) or a control group ( $n = 198$ ). Children were compared on measures of prosocial behavior, antisocial behavior, and related outcome measures. The trial found adverse effects on four outcomes among the intervention group compared to the control group: antisocial behavior increased on two different measures ( $d = 0.20$ ) and ( $d = 0.18$ ), child-reported liberal parenting increased ( $d = 0.16$ ), and child reported authoritarian parenting also increased ( $d = 0.20$ ). In addition, parental participation was significantly associated with several program outcomes. It was concluded, that group based after-school behavior programs may have the potential to cause iatrogenic effects and must be designed, piloted, evaluated and implemented with a high degree of care.

Liam O'Hare  
Andy Biggart  
Karen Kerr  
Paul Connolly

QUEEN'S UNIVERSITY  
BELFAST

---

**I**N recent times, the importance of children's social, personal, and behavioral outcomes has become prevalent in the international literature and has also received increasing attention within educational curricula. As a result a wide range of manualized programs have been designed to improve behavior and promote social skills. These programs are often referred to as social and emotional learning

programs (SEL), social skills training, or behavioral support programs. They are widely used, and some better known examples include PATHS (Promoting Alternative THinking Strategies; Greenberg, Kusche, Cook, & Quamma 1995), The Incredible Years (Webster-Stratton & Reid, 2003), Life Skills Training (Botvin & Griffin, 2002), and the Olweus Bullying Prevention Program (Olweus, Limber, & Mihalic, 1999).

The majority of SEL programs are delivered “in-school”; however, after-school programs have become increasingly popular. This is due to general concern over children’s outcomes (academic and social), a desire to improve community outcomes, and the provision of safe care for children while carers are at work (Little, Wimer, & Weiss, 2008). As both SEL programs and after-school provision have become increasingly popular, these two trends have converged and there has been a burgeoning of after-school programs that include elements focused on improving personal and social outcomes. The evidence behind the effectiveness of SEL programs is growing, but the majority of this research has been focused on in-school provision and there have been few rigorous evaluations of SEL programs in after-school settings. In order to add to the limited work in this area, this article focuses on presenting findings of a Randomized Controlled Trial (RCT) evaluation of an after-school prosocial behavior program. The study tested a number of research hypotheses. The primary hypothesis was that the Mate-Tricks after-school program would lead to moderate improvement in children’s behavioral outcomes in comparison to their control group peers. Other secondary hypotheses tested the impact of the program on related outcomes such as parenting, emotional intelligence, school attendance, and participation in clubs. Exploratory analyses examined differential responses to the intervention according to demographic factors, including gender, family affluence, and special educational needs. Further, exploratory analysis looked at several program implementation factors, including the number of sessions attended by the children and their parents as well as year cohort.

### **Research on the Effectiveness of Social and Emotional Learning Programs**

In order to get an overall picture of the evidence of SEL program effectiveness, Table 1 presents the combined results from a number of meta-analytic reviews that have investigated interventions to improve children’s personal and social outcomes (Diekstra, 2008; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; January, Casey, & Paulson, 2011; Lösel & Beelmann, 2003; Quinn, Kavale, Mathur, Rutherford, & Forness, 1999; Reddy, Newman, De Thomas, & Chun, 2009; Wilson, Lipsey, & Derzon, 2003). The study outcomes are separated into those that report immediate postintervention effects and those that report follow-up effects after a period of time. On closer inspection, the table highlights a number of salient points. It shows that most reviews report small to medium effects on a range of personal and social outcomes both in the short term (1,124 study outcomes and weighted mean effect size  $Md = 0.29$ ) and medium- to long-term outcomes (120 study outcomes and weighted mean effect size  $Md = 0.22$ ). So when comparing across the short- and longer-term outcomes, it would appear that the positive effects of these interventions persist, but decrease slightly over time. Furthermore, the reviews with the greatest number of study outcomes included in the meta-analysis consistently report smaller effects

Table 1. Effect Sizes from a Range of Meta-Analyses on SEL Interventions for a Range of Personal and Social Outcomes

Author	Outcomes	No. of Study Outcomes	Mean Effect Size (Cohen's <i>d</i> ) <sup>a</sup>
Diekstra (2008)	Academic achievement	9	.50
	Antisocial behavior	31	.48
	Clinical mental health problems	10	.16
	Positive self-image	6	.69
	Prosocial behavior	6	.59
	Social skills	31	.74
	Substance misuse	10	.11
Lösel & Beelmann (2003)	Antisocial behavior	80	.26
	Social skills	61	.39
	Social cognitive skills	57	.40
Quinn et al. (1999)	Mental disorders	35	.20
Reddy et al. (2009)	Emotional disturbance	5	.42
Wilson, et al. (2003)	Aggressive behavior	334	.23
January et al. (2011)	Social skills	28	.15
Durlak et al. (2011)	SEL skills	68	.57
	Attitudes	106	.23
	Positive social behavior	86	.24
	Conduct problems	112	.22
	Emotional distress	49	.24
Weighted mean on short-term outcomes <sup>b</sup>		1,124	.29
Diekstra (2008)	Academic achievement	7	.25
	Antisocial behavior	14	.17
	Mental disorders	8	.37
	Positive self-image	9	.08
	Prosocial behavior	6	.13
	Social skills	13	.05
	Substance misuse	15	.20
Lösel & Beelmann (2003)	Antisocial behavior	20	.22
	Social skills	16	.38
	Social cognitive skills	12	.33
Weighted mean on mid- to long-term outcomes <sup>b</sup>		120	.22

<sup>a</sup> Occasionally, reviews report negative mean effect sizes (–) which refer to a reduction in negative behavior. For comparison purposes, all effects sizes reported here are in relation to a positive change in behavior rather than actual direction of change given by the measure.

<sup>b</sup> This is calculated from (number of study outcomes × effect size)/total number of study outcomes.

(below  $d = 0.3$ ). Therefore, SEL interventions would be expected to have small to moderate effects on outcomes. Another emerging finding is that study outcome measures tend to report larger effects for increasing positive outcomes (mean effect size  $Md = 0.36$ ) than for reducing negative outcomes (mean effect size  $Md = 0.25$ ), although this difference is not significant in the small sample of meta-analyses reported in Table 1 (as assessed by a nonparametric  $t$ -test). Overall, small effect sizes appear common for SEL interventions. One potential reason for this is that some of these studies have had no effect or even caused a negative effect. In fact, it is estimated that up to 29% of these programs can have adverse effects (Lipsey, 1992). In essence, personal and social outcomes have the potential to improve but also decline depending on the impact of the program.

As the current study investigates an after-school program, it is important to look at the evidence of program effectiveness in this specific context. A relatively small

number of studies have investigated the impact of after-school programs on student outcomes across a range of dimensions, that is, developmental, academic, behavioral, social, and emotional (Gottfredson, Cross, Wilson, Rorie, & Connell, 2010; Lauver, 2002; U.S. Department of Education, 2003; Weisman, Soule, Womer, & Gottfredson, 2001). A review of after-school programs by Zief, Lauver, and Maynard (2006) showed that 84% of the 92 outcomes evaluated (from five studies that met the inclusion criteria<sup>1</sup>) showed no significant differences between the intervention and control groups. Within these reviews, personal and social outcomes were occasionally assessed, and some significant adverse program effects were found. Furthermore, recent large, high-quality studies of after-school programs by James-Burdumy, Dynarski, and Deke (2008) and Linden, Herrera, and Grossman (2011) also found some negative effects on child-reported personal and social outcomes.

One recent major meta-analysis investigating after-school SEL programs, however, presented evidence to the contrary. Durlak, Weissberg, and Pachan (2010), in an analysis of 75 reports evaluating 69 different programs, found a significant positive effect of after-school programs on a range of personal, social, and academic outcomes (e.g., feelings and attitudes, behavioral adjustment, and school performance). Furthermore, they provided evidence that the effectiveness of these programs is moderated by a set of practices conducive to personal and social skill development, referred to as “SAFE” practices (Durlak et al., 2010, p. 295). The SAFE practices were derived from a range of recommendations made by researchers working on personal and social skill development in school and clinical settings. Specifically, the SAFE practice model is composed of the following four elements: Sequenced and step-by-step training approach (building up concepts sequentially), Active forms of learning (like group work and role plays), Focused time on skill development (sufficient time for each targeted skill to be developed), and Explicit learning goals (so children and young people know what they are expected to learn). Overall, the meta-analysis by Durlak et al. (2010) concluded that programs which followed SAFE practices were significantly more effective than those that did not feature these practices. In summary, it appears that after-school programs are a challenging context in which to improve personal and social outcomes. This is evidenced by prior studies with null or negative effects (James-Burdumy et al., 2008; Zief et al., 2006), reduced effects of non-SAFE programs (Durlak et al., 2010), and an identified disparity between the effectiveness of SEL programs delivered in and out of school (Durlak et al., 2011).

Research has suggested many reasons why particular SEL programs appear to be more effective than others. These can be summarized into two main themes: participant characteristics and implementation quality. Several participant-related factors are well researched and have been found to correlate with social learning program effects. These factors include the children’s age, gender, level of disadvantage, and level of risk. Meta-analyses by January et al. (2011) and Fossum, Handegard, Martinussen, and Morch (2008) reported that increasing age led to a reduced positive effect on the development of social skills, whereby preschool children and children in kindergarten seemed to benefit more than their primary- and secondary-level counterparts. Adolescents showed lower, but still significant improvements. This evidence suggests that early intervention is most effective when delivering SEL interventions. January et al. (2011) also showed that the effect of socioeconomic status on the outcomes of a program failed to reach statistical significance, yet the trend indicated that students coming from lower socioeconomic backgrounds gained more

than their middle-class peers. Dolan et al. (1993) also reported an “at-risk” effect (i.e., programs for children already exhibiting antisocial behavior) where only boys at high risk of behavior problems benefited from the two classroom-based interventions designed to reduce aggression and improve academic achievement. Several other researchers have shown that targeted programs with children at risk of behavior problems had bigger effects than universal programs (Quinn et al., 1999; Reddy et al., 2009; Wilson et al., 2003). However, targeted SEL programs with at-risk children and their families begin with the development of more basic social skills. They also tend to be more intense (i.e., they last longer and sessions occur more frequently) than universal SEL interventions and, as a result, are likely to cost more.

It is widely accepted that the effectiveness of SEL programs depends on the quality of implementation (Cross, Gottfredson, Wilson, Rorie, & Connell, 2010; Derzon, Sale, Springer, & Brounstein, 2005; Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005; Farrell, Meyer, Kung, & Sullivan, 2001; Gottfredson & Gottfredson, 2002). Furthermore, meta-analyses have provided evidence that the outcomes of programs linked to personal and social learning are limited by implementation quality (Durlak & DuPre, 2008; Durlak et al., 2010, 2011). Implementation quality has a number of important aspects, and the following six key factors have been identified from the literature: program implementation practices, program fidelity, appropriate program duration, sufficient implementation time, community implementation support and training, and parental engagement.

The first implementation quality factor, program implementation practices (such as the aforementioned SAFE practices), highlights the importance of a program being sequenced, active, focused, and explicit. Also, Lösel and Beelmann (2003) showed that active instructional modalities (i.e., group work such as role plays) demonstrated a significantly greater effect size than passive, traditional instructional modalities (i.e., presentations). The second indicator of implementation quality is fidelity to a program manual and its aims. Hallfors and Godette (2002) reported that only 19% of the evaluations of evidence-based prevention programs report fidelity, that is, whether or not the program is conducted in a consistent way with all participants. Dumas, Lynch, Laughlin, Smith, and Prinz (2001) stated that low fidelity would have a direct effect on the program’s validity and indirectly affect evaluation studies. The third factor of appropriate program duration is important with respect to the amount of exposure children have to the program, that is, a specific dosage of a program may be required to produce the desired effects (Durlak et al., 2010). The fourth factor surrounds implementation time or how long the program has been active. Sufficient implementation time may be required to allow the program time to embed (Adelman & Taylor, 2003; Dusenbury & Hansen, 2004). Thus, positive effects may not immediately be apparent in the first cohorts of children. The fifth factor is the importance of a community’s delivery and support systems (Cross et al., 2010). For example, the development of systems and practices whereby the implementation of the program can be monitored on an ongoing basis (Domitrovich et al., 2008) as well as training support for those who deliver the program (Lochman, Boxmeyer, Powell, Wells, & Windle, 2009; Thaker et al., 2008). The final major indicator of implementation quality is parental engagement. Generally, studies report that the association between parenting and child behavior is complex, with gender as well as the children’s age mediating the association (Rothbaum & Weisz, 1994). However, some evidence suggests that parents without appropriate parenting skills can lead to



poor child behavioral outcomes (Hutchings, 2012). With regard to programs that attempt to improve parenting, the age of the child has again been highlighted as influencing effectiveness (McCart, Priester, Davis, & Azen, 2006). Furthermore, Davis, McDonald, and Axford (2012) suggest that in order to be effective, programs with a parenting element need to be socially inclusive, regardless of whether or not the program is deemed “high quality.” In other words, in order to benefit an entire community, effective programs must be delivered to as many children and families as possible. In their paper, they suggest four key principles for doing this: “improving recruitment and retention; co-producing a programme so it is culturally sensitive; building social capital amongst service users, implement with the wider community; and ensuring programmes are sustainable” (Davis et al., 2012, p. 20).

Despite the stated potential importance of parenting in relation to children’s personal and social development, the influence of parenting on SEL program effects is an under-researched area in schools (Durlak et al., 2011). Furthermore, it is suggested that there is less research, if any, that has been conducted into the relationship between parenting and SEL program effects implemented in after-school settings.

Several other factors have been suggested that influence SEL programs specifically in the after-school context. James-Burdumy et al. (2008) argued that after-school behavior programs may be influenced by increased fatigue at the end of the school day that may cause “acting-out” behaviors from the children, different disciplinary standards between in-school and after-school contexts, and the potentially negative influences of a child’s peers attending a referral-based program. Little et al. (2008) highlight three factors that they argue are crucial for success in after-school programs. They also encapsulate many of the general issues with SEL program implementation quality highlighted above: “Access to and sustained participation in programs; quality programming particularly, appropriate supervision and structure, well-prepared staff, intentional programming; and partnerships with families, other community organizations, and schools” (Little et al., 2008, p. 6).

In summary, most interventions that have been designed to promote personal and social outcomes in young people show low to medium effectiveness. In addition, there is some evidence to suggest that these interventions can cause negative effects. There is a range of evidence on the influences of demographic variables and implementation quality on program effectiveness. Therefore, standard implementation models available, such as SAFE training practices (Durlak et al., 2010), can prove useful. Ultimately, when combining the potential for adverse program effects with a wide range of possible participant and implementation influences, a rigorous and robust evaluation is critical. In particular, such an evaluation should be an inherent component when this type of program is being piloted. This is particularly important given the impact of poor evaluation research designs, which are consistently linked with spurious or inflated effect sizes. Therefore, it is important that SEL program evaluations have robust research designs that ideally include a control group, randomized assignment of participants, and samples of sufficient size to detect statistical significance if it is present. It is also important that outcomes are clearly stated at the outset and measured in an unbiased way. In addition, information about how implementation processes and program content are related to program effects is very valuable. If these evaluation criteria are fulfilled, then it will help promote beneficial interventions and identify programs that have detrimental effects.

## The Mate-Tricks Program

The Mate-Tricks program is a manualized and sequenced after-school program designed to improve personal development and social outcomes among children ages 9 to 10 years (Irish fourth class). The current study evaluated its impact in a community in Dublin, Ireland, that has been designated as an area of particular socioeconomic disadvantage with high levels of unemployment. The area has a population of 23,312 residents, and over a third of these are under 15 years of age.

The program is a blend of two existing prosocial behavior programs: the Strengthening Families Program (SFP) (Kumpfer & Alvarado, 2003) and the Coping Power Program (CPP) (Lochman & Wells, 2002a). The programs were blended by choosing specific activities from each and combining them into a unique sequence of activities to improve personal and social outcomes. The SFP utilizes a social ecology framework of child development (Kumpfer & Turner, 1991). Social ecological models, such as those offered by Bronfenbrenner, highlight the importance of holistic environmental factors in a child's development, including peers, schools, parents, and the wider society and culture (Bronfenbrenner & Bronfenbrenner, 1979). Consequently, the SFP includes numerous activities that involve participant children and parents developing several aspects of their personal and social lives. The CPP is based on a social cognitive model (Lochman & Wells, 2002a). Social cognitive models, and related theories such as social learning theory, highlight the importance of an individual's observations on the development of their knowledge, attitudes, behaviors, and thoughts. The CPP includes activities to help children and parents model appropriate personal and social skills through strategies demonstrated by their tutors. As a result, implementation of the program places a strong emphasis on the training of the CPP providers/counselors (Lochman et al., 2009).

Looking at the previous evidence of effectiveness for the two programs, the SFP is placed as a "promising program" within the internationally recognized list of Blueprints Programs for Violence Prevention (Center for the Study of Violence Prevention, Blueprints, 2011). Promising programs are in the second tier of programs, while the first-tier interventions are considered "model programs" and have the best available evidence to support their effectiveness. A range of studies have reported on the effectiveness of the SFP. A quasi-experimental 5-year study using a retrospective pre/posttest parental questionnaire found medium to high effects for child-only, parent-only, and child-parent outcomes across three different age groups (Kumpfer, Greene, Whiteside, & Allen, 2010). The strongest effects were found for children between the ages of 6 and 11 years old. The program has shown promising results in different cultural contexts (Erikson, 2002) and outside the United States (Bool & Onrust, 2009; Orte, Touza, Ballester, & March, 2008). So, overall, these studies appear to provide evidence for positive effects of the SFP. However, a number of more rigorous studies using randomly assigned control group designs have found weaker effects. For example, Semeniuk et al. (2010) showed that the program had mixed effects on outcomes. As anticipated, youth hostility decreased, but there were negative effects on two outcomes: parent hostility increased, and parent positive problem solving decreased. Another randomized trial has shown that children receiving full exposure to the program evidenced moderate improvement effects, but those receiving a partial exposure (not all sessions) showed a moderate decline (Riesch et al., 2012). Gottfredson et al. (2006) also identified a significant adverse effect of the



program on children's reports of negative peer associations. In this study, there were major challenges with the recruitment and retention of participants, and these issues were suggested as a reason for this adverse effect. In addition, one cluster randomized trial study of an international translation (i.e., moving it to a foreign context) showed no discernible effects on outcomes (Skärstrand, Sundell, & Andréasson, 2010).

Several studies of the CPP have shown that the program has promising effects. A random assignment study by Lochman and Wells (2002b) showed significant effects, including reductions in boys' self-reported covert delinquency, reduction in parent-reported substance use (in their parent-child intervention), and improvement in the boys' school behavior for both the child-only and parent-child conditions. The CPP has also been shown to have effects at one-year follow-up, and there is evidence of "radiance effects," that is, substance misuse reduction for children in the same classroom as those who have received the CPP (Lochman & Wells, 2004).

In practical terms, Mate-Tricks is a multisession after-school program that ran throughout one school year. Mate-Tricks comprised 59 child, six parent, and three family sessions. Three cohorts of children and their parents participated in the program between 2008 and 2011. The activities in the Mate-Tricks program varied across the child-only, parent-only, and parent/child sessions, but all sessions lasted one and a half hours, each of which was broken down into several 10–20-minute activities. The 59 child sessions targeted communication, staying calm, problem solving, managing emotions, and self-awareness (taken from the SFP), and, in addition, awareness of feelings, self-control, coping, perspective taking, social problem solving and handling conflict, and peer pressure (taken from the CPP). Each session also had common core activities, that is, a snack time, an opening game (chosen from an index of games), a review of the previous session, and a closing game. The six parent-only sessions were based around the following themes: understanding your child, encouraging good behavior, limit setting, communication, and problem solving. There was also time for refreshments, a review of the child sessions to date, and personal social/emotional skill development activities (e.g., stress management) and parenting strategies (e.g., holding family meetings). The family element utilized the following themes: communication, family values, empathy, and problem solving. These sessions had games, demonstrations, and discussions about how they spend time together as a family (called "Our Time"). There was also opportunity to practice the things they had learned in their respective individual sessions (e.g., a role play of a family meeting). As Mate-Tricks is a blend of the activities from two other programs, it therefore adopts a theory of development that includes both social cognitive and social ecological models. The intended outcomes of this program, as stated in the Mate-Tricks manual, are to "Enhance children's pro-social development; Reduce children's anti-social behavior; Develop children's confidence and self-esteem; Improve children's problem solving skills; Improve child-peer interactions; Develop reasoning and empathy skills; Improve parenting skills; and Improve parent/child interactions" (CDI, 2010, p. 45).

Reflecting on these outcomes and how they overlap with the two main underpinning theories of Mate-Tricks (outlined above), it can be seen that there are number of outcomes that are targeted by a social cognitive approach (e.g., improving children's problem solving skills through demonstrations from the facilitator of problem-solving strategies) and a social ecological framework (e.g., improving peer/parent/child/sibling interactions by working together on resolving issues).

Finally, the Mate-Tricks manual places a strong emphasis on training, and explicitly outlines the training requirements and ongoing staff support (CDI, 2010). The manual outlines that training should provide a theoretical framework, practical application, and sessions that include a combination of participation as well as reflection and sharing. It is stated that staff must take part in training on the SFP (2 days) and CPP (2 days). There should also be training on the use of the Mate-Tricks manual, as well as training on first aid and health and safety. It is also recommended that training be made available in the area of anti-bias intercultural approaches for working with migrant children and their parents. Ongoing training takes the form of booster sessions, including “Communities of Practice” meetings (COPs), where all stakeholders meet to reflect on Mate-Tricks implementation issues.

## Method

### Design

The study utilized a 3-year rolling cohort design. In other words, the effectiveness of the intervention was not assessed until all three cohorts had completed the intervention. The results reported in this article are pooled from three successive cohorts evaluated between 2008 and 2011. A rolling cohort design ensured that the implementation of Mate-Tricks is looked at in a longitudinal way (i.e., over the three cohorts) and provided a sufficient sample size (guided by effect size calculations) for statistical comparisons. With the rolling cohort design, particular care was taken over the potential evaluation contamination effects (i.e., influencing effects through evaluation methodology) through the release of interim results. Interim results were withheld, as this may have had undesirable and unintentional effects on the delivery of the program and/or undermined the validity of the evaluation. Also, any interim outcomes would only be tentative or perhaps even misleading. Therefore, the main study outcome analysis was not completed until all cohorts had completed the program and the full set of outcome data were collected.

### Participants

The participants in the study were 9- and 10-year-old children from a total of eight different primary schools. Children were entered into the evaluation based on referrals from teachers. All referrers were informed about the following: the nature of the program; the intention to refer children from the “full continuum of need” (CDI, 2010, p. 398), that is, that it was a universal intervention for all types of children; how the random allocation of places works; and how children and their families would be notified of their allocation. The intervention children were allocated to the Mate-Tricks program, and the control group received no alternative program or service. With regard to the control group, there are limited alternative after-school options in the local area, and Mate-Tricks was the only large organized after-school program available. So the majority of control group children would have returned home after school to be cared for by parents or other extended family members. Table 2 summarizes the main characteristics of the sample and is broken down by intervention and control group in relation to gender, assessed special educational need (SEN), and cohort. No data were collected on the children’s ethnicity, but a concurrent evaluation of an after-school literacy program for younger children (5–6-year-olds) in the

Table 2. Breakdown of the Sample by Gender, Special Educational Needs, and Cohort

	Intervention Group		Control Group	
	<i>N</i> <sup>a</sup>	%	<i>N</i> <sup>a</sup>	%
Gender:				
Boys	163	54.0	158	55.4
Girls	139	46.0	120	44.6
Total	302	100.0	278	100.0
Special educational need:				
Yes	20	13.0	28	17.8
No	126	81.8	123	78.3
Don't know	8	5.2	6	3.8
Total	154	100.0	157	100.0
Cohort:				
Cohort 1	102	33.6	87	30.5
Cohort 2	101	33.2	100	35.1
Cohort 3	101	33.2	98	34.4
Total	304	100.0	285	100.0

<sup>a</sup>There are differences between the total *N*'s in Table 2 between demographic variables because the data were collected by different methods. The research team collected data on the cohort, children provided data on their gender, and teachers provided data on SENs. The total *N*'s on SEN are much lower than gender and cohort in this table because the teachers' response rates were lower than that of the children.

same neighborhood reported a population of 16% minority ethnic children in the local schools (Biggart, Kerr, O'Hare, & Connolly, 2013). While these children came from a wide variety of nationalities, the most common were African (in particular, Nigerian), Irish Travellers, and Eastern European migrants.

## Ethics

A statement of ethics was approved by the Research Ethics Committee of the School of Education at Queen's University Belfast, ensuring that the study complied with the ethical standards of the American Education Research Association (2011) and the British Educational Research Association (2004). It covered issues relating to consent, privacy, confidentiality and data storage, well-being and safety of participants, and the intellectual property rights of participants, as well as the wider ethical issues relating to research with children. All fieldworkers and project staff were police checked. All program referrals were made in consultation with the children's parents/carers and written informed consent was gathered. Service providers took responsibility for the collection of consent forms and associated information. As part of the consent process, parents consented to their child's participation in the program as well as to his/her participation in the evaluation. All control group children were offered a place on a 1-week behavior-focused summer school once all aspects of the evaluation had been completed.

## Randomization

There were seven after-school settings, each with places for 15 children. In total, 30 children were allocated to each setting, with half the children randomly allocated to the intervention group and the other half allocated to the control group. Parental

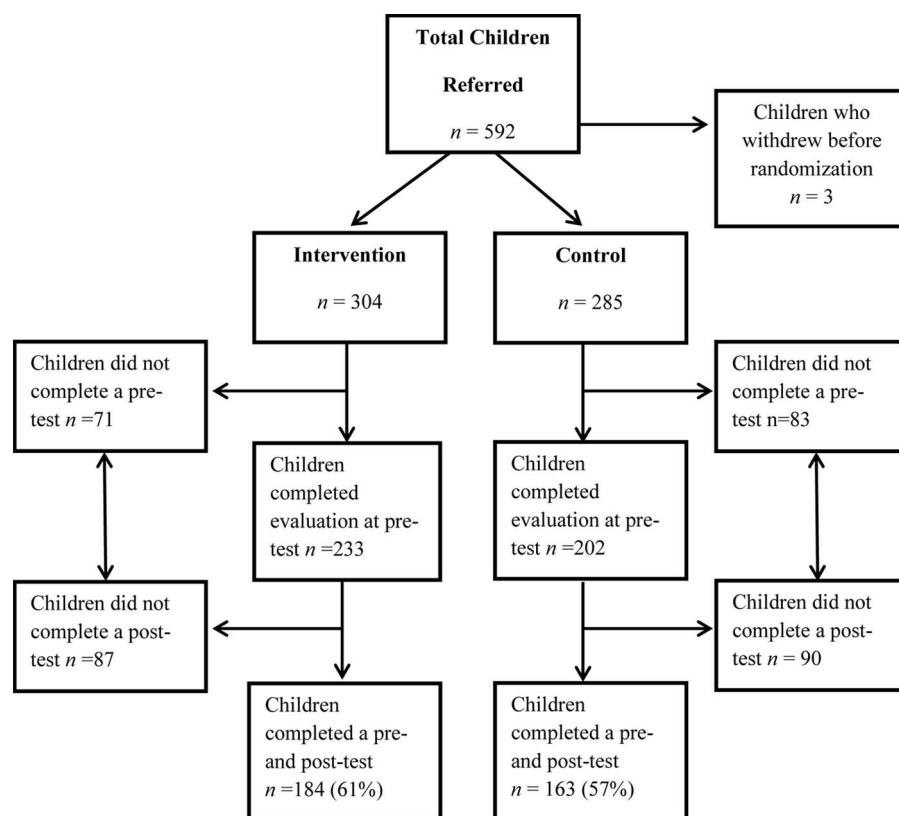


Figure 1. Participant flow through the study.

consent was sought prior to randomization and children were individually randomized within schools. For reasons of community transparency, the randomization process involved each child being allocated a raffle ticket that was placed in a hat and selected out by an independent observer who was a professional from the local education authority. After randomization, Cohorts 2 and 3 were closely matched in number. Cohort 2 had 101 intervention and 100 control children and Cohort 3 had 101 intervention and 98 control children. However, there was some disparity between the numbers in Cohort 1, with 102 in the intervention group and 87 in the control group. The reason for this was that, after randomization, it was discovered that a substantial number of siblings and twins had been allocated to different groups. This was undesirable for the service providers and the children's families. To rectify the situation, the twins or siblings were then allocated to the same group. Furthermore, it was decided that it was not desirable to remove children who had already been referred to the intervention group into the control group; therefore, all applicable children were moved into the intervention group, which resulted in the higher proportion in the intervention group in the first year. In Cohort 2 and Cohort 3, twins or siblings were allocated as a pair. Figure 1 highlights that there was some attrition during the various stages of testing. The major reason for this was that children were absent from school on the day of testing<sup>3</sup> or had left the school after referral was made. Overall, the pattern of attrition was similar in both the intervention and control groups, with similar numbers leaving the study at the various stages. In terms of teacher assessments, a similar response rate was obtained for those completing both pre- and

posttest (46% intervention  $n = 141$ ; and 49% control  $n = 138$ ). The main reason for attrition was that some teachers did not wish to provide information for the study or did not feel they had the time to do so.

### Assessment Procedures

In Mate-Tricks, children completed a pretest at the beginning of the program in September and the posttest took place at the end of the program in June. At both pre- and posttest, children in the intervention and control groups were tested in school during the day. The child testing was conducted in a group setting overseen by trained fieldworkers who were postgraduate students from a local university. All fieldworkers were provided with training in the standardized assessment procedures before they conducted the testing and followed the procedures as set out in the questionnaire administration manual. There was only one testing session for each school in order not to overburden schools with testing. Other teacher assessments in relation to child outcomes were undertaken by the regular class teacher. They completed a questionnaire for each child in their class, for both the intervention and control group. While it was impossible to blind participants to their involvement, the fieldworkers who collected outcome data were not informed of whether children were in the treatment or control groups.

### Study Measures

In the study, each of the outcome variables was measured by a composite mean score from several items within the research instrument. Outcome measures in the research instrument were chosen to map onto the theoretical frameworks and main outcomes of the program. Therefore, the research instrument measured both social cognitive (prosocial behaviors, attitudes, and skills) and social ecological constructs (parenting and peer relations). The outcomes in the study were measured using 16 different scales. Before the evaluation began, it was necessary to identify, a priori, a small number of primary study outcomes that were deemed to be the most important outcomes of the program. The remaining outcomes were regarded as secondary outcomes, and while still considered important, they were mainly conceived as supporting change in the primary outcomes. There were two primary outcomes: prosocial behavior and antisocial behavior. Prosocial behavior was assessed by a child self-assessment using a subscale of the Peer Relations and Prosocial Behavior Questionnaire (Rigby & Slee, 1992,  $\alpha = .75$ ). Antisocial behavior was measured in three ways. There were two subtly different child-reported measures that varied in the scope and language of question items. They were a subscale of the Peer Relations and Prosocial Behavior Questionnaire (Rigby & Slee, 1992,  $\alpha = .83$ ) and the children's version of the Child Behavior Checklist (Achenbach et al., 2008; this scale was adapted from a three-point Likert scale to a five-point scale to improve response variation,  $\alpha = .91$ ). There was also a teacher report on antisocial behavior using the adult version of the Child Behavior Checklist (Achenbach et al., 2008; again, the scale was adapted to use a five-point Likert scale,  $\alpha = .96$ ).

In addition, 12 secondary outcomes were measured. These outcomes included behaviors related to attention deficit hyperactivity disorder (ADHD), which were measured by a teachers' assessment using the ADHD Rating Scale (DuPaul, 1991,  $\alpha =$

.95). Victimization was assessed using children's responses on a subscale of the Peer Relations and Prosocial Behavior Questionnaire (Rigby & Slee, 1992,  $\alpha = .85$ ). Another instrument, the child-report version of the Alabama Parenting Questionnaire (Frick, 1991), was used. This scale was amended because some community stakeholders indicated that they were concerned about a few of the items and asked if these could be removed. Generally, they requested the removal of items that referred to parental corporal punishment. As a result, the scale underwent psychometric refinement of subscales (using factor analysis, which is not reported in this article due to space limitations). The resultant subscales were named as follows: Maternal Relationship ( $\alpha = .78$ ), Paternal Relationship ( $\alpha = .85$ ), Supportive Parenting ( $\alpha = .81$ ), Liberal Parenting ( $\alpha = .79$ ), and Authoritarian Parenting ( $\alpha = .71$ )<sup>2</sup>. Trait emotional intelligence was also assessed in two ways using the Trait Emotional Intelligence Questionnaire (Mavroveli, Petrides, Rieffe, & Bakker, 2007) through a child self-report ( $\alpha = .83$ ) and a teacher report ( $\alpha = .97$ ). Overall, the measures were found to show good reliability. However, one three-item measure of child-reported conflict tactics was removed from the analysis because it had low reliability ( $\alpha = .48$ ). Additionally, school attendance was retrieved from school records and a proxy measure of social engagement was measured by asking children how many clubs or societies they belonged to. A number of parental measures were also collected through a postal questionnaire designed to measure parental perceptions of their child's behavior and their relationship. However, despite significant and ongoing attempts to improve parental response rates, including reminder letters, questionnaire reissues, and telephone follow-up, the parental response rate remained below an acceptable level (17% completing both pre- and posttest), and the results are not reported in detail in this report.<sup>4</sup> This highlights the serious difficulties in securing high response rates from parents, and while there could be a number of potential barriers to completion, some may have had significant literacy difficulties themselves. Telephone completion was specifically offered to mitigate this issue, although many did not avail themselves of this.

### Pretest Scores and Study Analysis

Table 3 provides the pretest scores for the control and intervention groups. The primary outcome pretest means for the control and intervention groups provide some indication of the general behavior of the children before their participation in the study. The measures' maximum and minimum scores are as follows: prosocial behavior PSBQ, 5 = most prosocial and 1 = least prosocial; antisocial behavior PSBQ (child report), 5 = most antisocial and 1 = least antisocial; antisocial behavior CBCL (child report), 5 = most antisocial and 1 = least antisocial; antisocial behavior CBCL (teacher report), 4 = most antisocial and 0 = least antisocial. Therefore, looking at Table 3, it can be seen that overall the general behavior of the sample was good with some room for improvement. The analysis of pretest scores in Table 3 also shows that the control and intervention groups were well matched. These results indicated that attrition rates had not led to any notable biases being introduced into the trial in terms of creating nonequivalent groups. Notably, a number of participants provided posttest data but no pretest data. In order to ensure that their posttest data were included in the outcome analysis, it was deemed appropriate to impute pretest data for these individuals. This led to an average of 26% of the pretest scores being im-



Table 3. Means, Standard Deviations, and Significance of Difference on Outcome Variables at Pretest for Children and Teachers

Variable	Mean Intervention (SD)	Mean Control (SD)	Sig.
Primary outcomes:			
Prosocial behavior PSBQ (child report)	4.02 (.93)	4.05 (.87)	$t(587) = .40, p = .69$
Antisocial behavior PSBQ (child report)	1.63 (.73)	1.71 (.84)	$t(587) = 1.29, p = .20$
Antisocial behavior CBCL (child report)	1.72 (.57)	1.79 (.64)	$t(587) = 1.35, p = .18$
Antisocial behavior CBCL (teacher report)	.51 (.61)	.48 (.57)	$t(587) = -.71, p = .48$
Secondary outcomes:			
Attendance at school	90.67 (12.09)	90.89 (10.73)	$t(395) = .20, p = .85$
ADHD-related behaviors (teacher report)	.75 (.58)	.74 (.54)	$t(587) = -.36, p = .70$
Victim perceptions (child report)	2.57 (1.05)	2.60 (1.01)	$t(587) = .30, p = .76$
Clubs attended (child report)	1.38 (1.47)	1.50 (1.48)	$t(435) = .83, p = .41$
No. of friends (child report)	3.26 (1.15)	3.26 (1.06)	$t(423) = .61, p = .54$
Maternal relationship (child report)	3.65 (.80)	3.66 (.79)	$t(587) = .11, p = .91$
Paternal relationship (child report)	3.34 (.97)	3.38 (.89)	$t(587) = .54, p = .59$
Liberal parenting (child report)	2.29 (.57)	2.38 (.61)	$t(587) = 2.00, p = .05$
Supportive parenting (child report)	3.98 (.83)	3.96 (.78)	$t(587) = -.40, p = .69$
Authoritarian parenting (child report)	2.56 (.88)	2.52 (.88)	$t(587) = -.48, p = .63$
Trait emotional intelligence (teacher report)	6.55 (1.47)	6.50 (1.55)	$t(587) = -.40, p = .69$
Trait emotional intelligence (child report)	3.48 (.44)	3.48 (.37)	$t(587) = -.03, p = .98$

puted (varying slightly for different items). The imputation method used was single imputation that utilized an EM algorithm (Dempster, Laird, & Rubin, 1977), and imputation was conducted on the underlying data (i.e., raw data, and not at the subscale or outcome level). The treatment-group variable, that is, control or intervention, was not included in the imputation process. As no actual outcome data (posttest data) were imputed, it was felt that single imputation was sufficient, as it would only make minor adjustments to the pretest covariate data, and the main impact of the imputation was that a greater number of children could be included in the outcome analysis and thus increase study power. The main statistical analysis of program effects was conducted using multiple linear regression. All primary outcome variables were included as covariates in the regression models in order to try to maximize statistical power. The sample was individually randomized, but the program was delivered across seven sites, so the standard errors in the regression analysis were adjusted for the seven clusters using robust standard errors in stata (Maas & Hox, 2004). Adjusted posttest means were also calculated for each of the groups,

controlling for pretest scores. Effect sizes were then calculated as standardized mean differences (Cohen's *d*). Further prespecified exploratory analyses of demographic variables were conducted by adding interaction terms into the regressions, along with the primary outcome covariates used in the main analysis models. This was done to examine subgroup effects for boys and girls, family affluence/poverty (as measured by the Family Affluence Scale II, see Schnohr et al., 2008; Kehoe and O'Hare, 2010), and special educational needs. In addition, analyses were conducted to explore whether or not several implementation variables were related to better outcomes for the intervention children. These variables included cohort year, the number of sessions attended by the child (i.e., the level of exposure to the program or dosage), and the number of sessions attended by the child's parent or guardian. Family affluence scores were included in the exploratory models of parent and child attendance to control for any variations in family background.

### Study Power

A sample-size power calculation was conducted based on identifying an effect size with  $d = 0.2-0.4$ , a statistical power level of 0.8, having a minimum of two predictors in the model, and identifying a  $p < .05$ . The desired sample size was calculated to be in the range of  $N = 241-478$ . It was proposed that 210 children would be referred to the Mate-Tricks program each year for 3 years, giving a maximum proposed total sample of 630 children. The initial child sample for the study at pretest was 592, which exceeded the sample size required. After attrition and imputation, the sample still remained within required sample-size range ( $N = 302-402$ ).

### Process Evaluation

In conjunction with the RCT, a process evaluation was conducted. This element of the study included data analysis of in-depth interviews, site observations, and documentation. The main purpose of this investigation was to ascertain how the program was being delivered across different sites, identifying any variations in implementation and any other relevant factors where differences may have been evident, for example, the number of children attending/dropping out, parental participation, timetable restrictions, and resources available. In addition, the process evaluation was used to provide further insights into which elements of the program tended to work or not, and the reasons why. The in-depth interviews were conducted with all facilitators, service provider staff, four principals, and three members of staff in the coordinating community organization (the implementation team). Focus groups were carried out with two groups of children and two groups of parents. Site observations were undertaken at all sites, and an analysis of documentation was conducted (e.g., minutes from the communities of practice meetings).

## Results

### Main Analysis

Table 4 presents the regression models for all the child and teacher outcome measures. Table 5 summarizes the main effects and reports: the adjusted posttest score (calculated from the unstandardized beta coefficients from the regressions) with

Table 4. Regression Coefficients B (with Robust Standard Error) of Mate-Tricks Program Effects Including  $R^2$  and Sample Size N

Variables	Intervent.	Pretest	PSBQ Pro	PSBQ Anti	CBCL Child	CBCL Teacher	Constant	$R^2$	N
Primary outcomes:									
Prosocial behavior PSBQ (child report)	-.068 (.117)	.316 (.087)	—	.017 (.045)	-.100 (.060)	-.205 (.189)	3.053 (.431)	.124	402
Antisocial behavior PSBQ (child report)	.150* (.053)	.185 (.056)	-.152 (.046)	—	.085 (.065)	.181 (.115)	1.545 (.216)	.160	396
Antisocial behavior CBCL (child report)	.106* (.044)	.274 (.063)	-.066 (.036)	-.015 (.035)	—	.137 (.025)	1.303 (.172)	.185	330
Antisocial behavior CBCL (teacher report)	-.006 (.070)	.569 (.156)	-.127 (.035)	-.047 (.045)	.107 (.034)	—	.571 (.214)	.406	351
Secondary outcomes:									
Attendance at school	-.260 (.606)	—	-.335 (.721)	-.117 (.700)	.770 (.498)	-.890 (.654)	93.321 (3.999)	.007	396
ADHD-related behaviors (teacher report)	-.036 (.088)	.673 (.052)	-.088 (.023)	.083 (.027)	-.008 (.052)	.156 (.056)	.423 (.133)	.517	377
Victim perceptions (child report)	-.007 (.096)	.528 (.060)	-.129 (.067)	-.005 (.098)	.454 (.073)	.031 (.098)	1.551 (.465)	.233	397
Clubs attended (child report)	.153 (.198)	.415 (.064)	-.126 (.101)	.136 (.144)	-.324 (.202)	.320 (.194)	1.933 (.529)	.154	346
No. of friends (child report)	-.048 (.102)	.440 (.057)	.041 (.067)	.067 (.042)	.061 (.090)	-.037 (.057)	1.504 (.529)	.212	303
Maternal relationship (child report)	.186 (.125)	.574 (.092)	-.077 (.035)	.149 (.058)	-.121 (.076)	-.057 (.031)	1.848 (.389)	.257	353
Paternal relationship (child report)	.166 (.107)	.579 (.088)	-.113 (.047)	.032 (.115)	-.033 (.152)	-.024 (.076)	1.867 (.291)	.256	323
Liberal parenting (child report)	.104* (.043)	.307 (.089)	-.124 (.056)	-.039 (.042)	.120 (.103)	-.048 (.054)	1.810 (.324)	.163	302
Supportive parenting (child report)	.108 (.104)	.544 (.042)	-.058 (.031)	.155 (.054)	-.130 (.043)	-.079 (.022)	2.105 (.249)	.255	357

*Continued.*

Table 4. (Continued)

Authoritarian parenting (child report)	.204 *	.411	.070	.091	-.084	-.062	1.179	.163	378
	(.062)	(.070)	(.060)	(.109)	(.064)	(.056)	(.341)		
Trait emotional intelligence (teacher)	-.074	.723	.248	.037	-.186	-.088	1.030	.499	378
	(.156)	(.045)	(.104)	(.083)	(.255)	(.114)	(.733)		
Trait emotional intelligence (child report)	-.057	.469	.068	.009	-.043	-.062	1.797	.276	418
	(.029)	(.034)	(.018)	(.023)	(.020)	(.062)	(.096)		

\*Indicates significant coefficient for "intervention" at the .05 level.

Table 5. Adjusted Posttest Scores of Measures for Intervention and Control Groups with Effect Sizes and Significance Values

Variable	Adjusted Posttest Scores <sup>a</sup> (with <i>SD</i> )		Effect Size ( <i>d</i> ) (95% Confidence Interval)	Significance <i>p</i>
	Intervention	Control		
Primary outcomes:				
Prosocial behavior PSBQ (child report)	4.01 (.35)	4.08 (.33)	−.07 [−.29, +.16]	.58
Antisocial behavior PSBQ <sup>b</sup> (child report)	1.63 (.31)	1.49 (.32)	.18 [+.06, +.31]	.03
Antisocial behavior CBCL <sup>b</sup> (child report)	1.66 (.21)	1.57 (.23)	.20 [+.04, +.36]	.05
Antisocial behavior CBCL <sup>b</sup> (teacher report)	.46 (.39)	.44 (.37)	−.01 [−.22, +.20]	.93
Secondary outcomes:				
Attendance at school	90.67 (.80)	90.90 (.86)	−.02 [−.13, +.08]	.69
ADHD-related behaviors <sup>b</sup> (teacher report)	.75 (.45)	.77 (.49)	−.05 [−.27, +.18]	.70
Victim perceptions (child report)	2.47 (.56)	2.49 (.53)	−.01 [−.16, +.15]	.95
Clubs attended (child report)	1.98 (.66)	1.85 (.67)	.09 [−.14, +.32]	.47
No. of friends (child report)	3.22 (.52)	3.31 (.48)	−.05 [−.23, +.14]	.66
Maternal relationship (child report)	3.83 (.44)	3.65 (.42)	.21 [−.07, +.48]	.19
Paternal relationship (child report)	3.49 (.55)	3.35 (.48)	.16 [−.04, +.36]	.17
Liberal parenting <sup>b</sup> (child report)	2.24 (.24)	2.17 (.26)	.16 [+.03, +.27]	.05
Supportive parenting (child report)	4.14 (.43)	4.02 (.41)	.12 [−.11, +.35]	.34
Authoritarian parenting <sup>b</sup> (child report)	2.68 (.37)	2.47 (.37)	.20 [+.08, +.32]	.02
Trait emotional intelligence <sup>c</sup> (teacher report)	6.39 (1.18)	6.43 (1.20)	−.04 [−.20, +.12]	.66
Trait emotional intelligence <sup>c</sup> (child report)	3.55 (.26)	3.61 (.22)	−.13 [−.24, +.00]	.10

<sup>a</sup> Controlling for pretest score (exception for attendance at school, collected at posttest only).<sup>b</sup> A higher score indicates a positive outcome, with the exception of these items where a higher score indicates an increase in problematic behaviors/attitudes.<sup>c</sup> Teacher and child ratings differ because the measures are scaled differently.

standard deviations, the effect size difference between the control and intervention group on each of the primary outcomes, the effect size confidence intervals, and statistical significance. With regard to the four measures of primary behavioral outcomes (Table 5), those children who participated in Mate-Tricks reported significantly higher levels of antisocial behavior than the control group children on two measures (PSBQ  $d = 0.18$  and CBCL  $d = 0.20$ ). No evidence was found of significant effects in relation to the other two primary outcome measures. With regard to the secondary outcomes, there were two significant effects from the 12 reported measures. The children who received Mate-Tricks reported significantly higher levels of

authoritarian parenting ( $d = 0.20$ ) and liberal parenting ( $d = 0.16$ ). None of the other secondary outcomes showed significant changes, including school attendance, ADHD-related behaviors, club participation, and trait emotional intelligence.

### Exploratory Analysis

The prespecified exploratory subgroup analyses were undertaken to see whether Mate-Tricks worked differently for a number of participant groups based on a range of demographic variables: gender, family affluence, and special educational needs. There were two significant interactions for gender on outcome variables, where intervention boys reported significant decreases compared to control group boys on authoritarian parenting ( $p = .01$ ) and relationship with mothers ( $p = .02$ ). There was only one significant interaction for family affluence, which was that being in the intervention group predicted a significant decrease in supportive parenting ( $p = .03$ ). There were no significant interactions between special educational needs and any program outcomes. In addition to these demographic variables, a further exploratory analysis was conducted to explore three implementation variables: the cohort the child participated in, the number of sessions attended by the child (i.e., their level of exposure to the program or dosage), and the number of sessions attended by the child's parent or guardian. There were a number of significant results identified in these exploratory analyses, and these are outlined in the paragraphs below.

**Cohort.** As outlined in the methodology, there were three cohorts involved in the study. Several significant cohort effects were found, depending on which year the families participated in the program. Specifically, teachers reported significantly increased levels of antisocial behavior (as measured by the CBCL) for intervention children in Cohorts 2 ( $p = .01$ ) and 3 ( $p = .01$ ) compared to Cohort 1. Teachers also reported increases in ADHD-related behaviors for Cohort 2 intervention children ( $p = .02$ ) compared to those in Cohort 1. The intervention had a positive effect on child-reported relationships with their mothers in Cohort 3 compared to Cohort 1 ( $p = .02$ ) and relationships with their fathers in Cohort 2 compared to Cohort 1 ( $p = .02$ ). There was also an increase in supportive parenting for Cohort 2 intervention children compared to Cohort 1 ( $p = .01$ ).

**Number of Mate-Tricks sessions attended by children.** It is important to note that this aspect of the exploratory analysis was restricted to children in the intervention group only. Therefore, the results are predictive rather than causal. It was designed to test whether the extent of children's exposure to the Mate-Tricks program had any predictive relationship with the outcome variables. Program facilitators had been asked to keep a weekly register of children who attended the program. The children in the intervention group received an average of 23.00 sessions ( $SD = 19.81$ ) that equated to an average of  $M = 34.5$  hours contact time per child. The minimum number of sessions that a child attended was 0 and the maximum number was 59. The analysis found that children's level of exposure to the program was not predictive of any of the program outcomes.

**Number of Mate-Tricks sessions attended by parents.** The same analysis procedure carried out on child exposure was also carried out on parental exposure. The parents in the intervention group received an average of 2.34 sessions ( $SD = 2.64$ ) that equated to an average of  $M = 3.51$  hours contact time per parent. The minimum number of sessions parents were involved in was 0 and the maximum number was 9.



Table 6. Regression Coefficients B (with Robust Standard Error) of the Effects of Parental Attendance on Program Outcomes Including  $R^2$  and Sample Size  $N$ 

Variable	Parent Attendance	Pretest Score	Family Affluence	Constant	$R^2$	$N$
Primary outcomes:						
Prosocial behavior PSBQ (child report)	.078 <sup>a</sup> (.019)	.392 (.126)	-.005 (.033)	2.251 (.635)	.168	214
Antisocial behavior PSBQ (child report)	-.044 (.021)	.404 (.082)	-.089 (.032)	1.546 (.131)	.138	206
Antisocial behavior CBCL (child report)	-.012 (.015)	.292 (.043)	-.017 (.021)	1.282 (.218)	.085	171
Antisocial behavior CBCL (teacher report)	.009 (.023)	.605 (.214)	-.024 (.050)	.251 (.342)	.343	178
Secondary outcomes:						
Attendance at school	.718 (.384)	—	.569 (.711)	85.893 (3.147)	.028	197
ADHD-related behaviors (teacher report)	-.026 (.021)	.830 (.052)	.034 (.034)	.017 (.244)	.542	196
Victim perceptions (child report)	.036 (.021)	.474 (.082)	-.049 (.058)	1.404 (.245)	.193	210
Clubs attended (child report)	-.041 (.078)	.369 (.070)	-.017 (.079)	1.662 (.616)	.124	183
No. of friends (child report)	-.022 (.017)	.503 (.057)	.002 (.030)	1.66 (.144)	.252	176
Maternal relationship (child report)	.034 (.018)	.533 (.078)	.011 (.043)	1.727 (.437)	.276	186
Paternal relationship (child report)	.066 <sup>a</sup> (.026)	.490 (.080)	<.001 (.071)	1.700 (.547)	.240	167
Liberal parenting (child report)	-.046 <sup>a</sup> (.019)	.402 (.145)	-.017 (.016)	1.52 (.417)	.165	156
Supportive parenting (child report)	.018 (.018)	.476 (.088)	-.031 (.048)	2.344 (.287)	.233	188
Authoritarian parenting (child report)	.061 <sup>a</sup> (.020)	.433 (.106)	.004 (.021)	1.393 (.317)	.184	197
Trait emotional intelligence (teacher report)	.071 (.055)	.763 (.067)	-.008 (.065)	1.252 (.583)	.495	196
Trait emotional intelligence (child report)	.033 <sup>a</sup> (.012)	.489 (.054)	.016 (.014)	1.687 (.156)	.277	220

<sup>a</sup>Indicates parent attendance coefficient significant at the .05 level.

This variable was found to be the most significant predictor of program outcomes. Parents/guardians attending sessions more often predicted a wide range of child outcomes: an increase in child-reported prosocial behavior ( $p = .01$ ), an improvement in child-reported relationships with their male guardian ( $p = .05$ ), an increase in authoritarian parenting ( $p = .02$ ), a reduction in liberal parenting ( $p = .05$ ), and an increase in child-reported trait emotional intelligence ( $p = .03$ ). As parent attendance had the greatest predictive power on outcomes, the full regression models for this analysis are provided in Table 6.

### Process Evaluation

The focus of this article is on the reporting of program effects, and thus it is not possible to fully report the process evaluation. However, the process evaluation is reported elsewhere (O'Hare, Kerr, Biggart, & Connolly, 2012). The key points to

draw out in terms of participants and implementation are as follows. Evidence from the interview data indicated that those who engaged with the Mate-Tricks program were positive about the impact of the program on the children involved and about the program itself. Interviewees were positive about specific aspects of either the content of the program and the approaches used to deliver it, to include the fact that it was a manualized approach. The facilitators were also positive about the planning, reflection processes, training, and use of a co-facilitation approach.

In terms of negative comments, interviewees raised issues related to behavior problems in Mate-Tricks and about how suitable Mate-Tricks was for “all types of children,” for example, children with behavioral problems. Parental attendance was consistently cited as an ongoing issue. However, it is important to note that the service providers made numerous changes, and put several strategies in place, in an attempt to reach and engage a greater number of parents, including increasing the overall number of sessions to facilitate group belonging, providing repeat or “catch-up” sessions, parent sessions taking place closer together (over a period of 6 weeks instead of being split up over the academic year), house visits that kept parents up to date on the sessions, communication by text and telephone, reorganization of the starting time of parent sessions, amalgamation of groups in order to offer two alternate session times, and general positive communication with parents when collecting children from the program. In short, improving parental engagement required substantial extra commitment and resources.

## Discussion

In relation to the primary research question—Does Mate-Tricks improve children’s behavioral outcomes?—the results do not provide evidence that the program was effective in improving this key outcome. In fact, there were two statistically significant, undesirable effects of the program on behavior, with an increase in child-reported antisocial behavior as measured by the PSBQ ( $d = 0.18$ ) and CBCL ( $d = 0.20$ ). There was also an iatrogenic effect on two secondary outcomes, with significantly higher levels of authoritarian parenting ( $d = 0.20$ ) and liberal parenting ( $d = 0.16$ ) reported by the children, both of which were deemed undesirable parenting practices. Although these four negative effects were not desirable, they are not unusual given the substantial minority of negative effects that have been found among previous rigorous evaluations of SEL programs (estimated to be at least 29%; Lipsey, 1992). Furthermore, several recent RCT studies have reported adverse effects of the SFP, which provided many of the activities that form part of the Mate-Tricks program (Gottfredson et al., 2006; Riesch et al., 2012; Semeniuk et al., 2010). More generally, this evaluation contributes to the wider body of evidence showing that interventions focused on changing psychological outcomes have the potential to cause adverse effects as well as benefits (Lilienfeld, 2007).

These findings also concur with some of the previous evidence of effectiveness in relation to after-school social learning programs. The current study showed no effects on the majority of outcomes. Similarly, the meta-analysis by Zief et al. (2006) showed that 84% of the outcomes measured in behavioral after-school programs showed no effects. Also, the resultant negative effects observed in the current study, are consistent with the findings presented in recent, large-scale studies of after-school programs (James-Burdumy et al., 2008; Linden et al., 2011). However, the

findings in this study deviate from the findings by Durlak et al. (2010) in relation to the effectiveness of SAFE social and emotional learning after-school programs as the Mate-Tricks program followed SAFE practices in that it was sequenced (builds on concepts over the sessions), involved active learning (e.g., role-plays), included a focus on skill development (such as conflict resolution and emotional skills), and the children, facilitators, and parents were explicitly made aware of the program outcomes. The Durlak et al. (2010) meta-analysis would suggest that following these practices can enhance effects, but in this instance following the SAFE model was not a guarantee of success. This would infer that some factors other than program practices are a limiting factor on the effects of this particular program.

With the current evaluation results in mind, it is useful to explore the additional factors that may be associated with the negative outcomes identified in the current study. If we examine which factors appear to be associated with the effectiveness (or iatrogenic effects) of the Mate-Tricks program, the exploratory analysis suggested that participant factors such as gender, special educational needs, and disadvantage (as measured by family affluence) had only a small impact on program effects (only three significant results from 64 models). Rather, the evidence produced in the current evaluation showed that implementation factors had greater predictive strength. The literature has suggested several implementation factors that may be associated with the effectiveness of after-school programs. These are adequately summarized by the three issues highlighted by Little et al. (2008): program quality, partnerships, and participation. With regard to program quality, it is argued that Mate-Tricks successfully delivered on many of the key aspects: there was appropriate supervision and structure provided by teachers and community workers, with well-prepared and trained facilitator staff, and there was intentional programming that reflected recommended SAFE practices. With regard to the second issue (partnerships), there were strong partnerships between schools, community organizations, and numerous parents during the 3-year pilot. However, the data suggest that it is the last issue outlined by Little et al. (2008)—engaged participation by parents—that appeared to be the most influential factor in this study. A greater number of Mate-Tricks sessions attended by parents predicted changes across many child outcomes, including increasing prosocial behavior, improving relationships with fathers, increasing authoritarian parenting, decreasing liberal parenting, and increasing trait emotional intelligence. This mixed pattern of predicting outcomes (mostly positive, but also an increase authoritarian parenting) shows a significant but complex relationship between parental attendance and the Mate-Tricks program outcomes. In addition, the process evaluation highlighted an overall lack of engaged participation from parents by the program facilitators. Overall, these findings cannot attribute causality, but they do highlight an important future research agenda exploring the complex influences between children, parents, and SEL program outcomes.

These exploratory findings may suggest that if Mate-Tricks were to be adapted and focused on the recruitment of parents and children who are likely to fully engage with the program, it could lead to positive effects. However, this is problematic because the recruitment of parents who will participate may be a difficult task. Furthermore, a program that serves only a subsection of the community, in this case engaged parents and children, would not meet the original aims of a program that was based on universal community need. Such an approach may also exclude the children with the most difficult behaviors and potentially increase inequality in

terms of outcomes. Furthermore, this would no longer constitute a “socially inclusive programme,” the importance of which is highlighted by Davis et al. (2012). The issue of developing a socially inclusive program reinforces the need for future research to identify the best ways to engage the participation of the “hardest to access” parents in a community of social disadvantage, particularly when implementing universal after-school behavior programs. There are few examples of universal out-of-school SEL programs like Mate-Tricks; however, one existing example is the Incredible Years “Attentive Parenting” Program. This has yet to undergo a rigorous evaluation with a control group present (unlike the other versions of the Incredible Years program suite). Furthermore, this program cannot be considered fully universal as it is not recommended for children and parents with high levels of difficulties (Webster-Stratton, 2012).

Reflecting on potential reasons for the negative effects identified in this study, several suggestions can be made. The first potential explanation is the counterfactual to the intervention; that is, the control group was exposed to situations more attuned to improving behavior (or at least maintaining current levels of behavior) than Mate-Tricks. There are limited after-school activities in the community under investigation, and Mate-Tricks was the only large-scale program operating in the area; so most control children were likely to be going home and cared for by their parents and/or extended family members, which possibly includes play with siblings and peers. This may provide a better environment for promoting social and emotional learning. However, without further information on the control group’s activities, it is not possible to discern the impact of the counterfactual in the current study.

Another explanation is that it has been hypothesized that evaluations detecting negative effects of after-school behavioral programs have simply measured an increase in children’s sensitivity to, or awareness of, their bad behavior, and as a result children report higher levels of antisocial behavior (Linden et al., 2011). However, there are several reasons why this may not be the case. The well-recognized effects of social desirability, the Hawthorn effect and the Pygmalion effect, would act in the opposite direction to this hypothesized sensitivity effect. In other words, children within the intervention group may have evidenced an increased desire to portray themselves in a better light because they were involved in the Mate-Tricks program. Second, the logic of a sensitivity effect would suggest that if children were more sensitive to their antisocial behavior then they would also have increased sensitivity to their prosocial behavior. However, the changes in prosocial behavior in this study were also in a negative direction (although not significant). Finally, given the consistency of the negative impact on antisocial outcomes in this study, it would be hard to justify the results as simply an increase in sensitivity.

One of the most widely cited explanations of iatrogenic effects found in the evaluation of SEL programs is the negative effects of peer influences (Dishion & Kavanagh, 2003; Piehler & Dishion, 2007). In essence, this research suggests that peers learn negative behaviors from each other when instructed in groups. This is particularly the case for referral based programs when groups of children with existing problem behaviors are clustered together. This could potentially explain the negative effects found in relation to Mate-Tricks on antisocial behavior. However, Mate-Tricks was a universal rather than selective referral program. Furthermore, the results from the study suggest a more complex picture, with negative effects on par-

enting as well as antisocial behavior (including associations between parental participation and outcomes).

This leads to one further hypothesis offered by the article's authors. This hypothesis is derived from an overview of the iatrogenic effects in this study on behavior and parenting as well as the significant relationships between parental attendance and program outcomes. The hypothesis is that the Mate-Tricks program may have created an imbalance between the parents and children regarding their personal and social knowledge. Hypothetically, this provided the conditions for competitive emotional learning between the children and their parents rather than cooperative emotional learning. For example, a child's newly acquired personal and social knowledge may have resulted in a shift in authority between the parent and child, with the parent "giving up" and becoming more liberal (particularly if the parent was not aware of the program content through their nonparticipation). Conversely, a parent may have attended the program and felt compelled to adapt their parenting style, but did so by simply becoming more authoritarian. Both of these hypothesized processes could have led to the observed negative effects on children's antisocial behavior either through the children's frustration with increased levels of authoritarian parenting, resulting in "acting out" behaviors, or increased opportunity for children to engage with antisocial behavior as a result of more liberal parenting. Ultimately, this hypothesis is one without supporting evidence in the existing SEL program evaluation literature. Therefore, it is a tentative hypothesis and one that would require substantial future research to establish its validity.

There are several limitations to consider in the study. The first is related to the use of multiple testing and the potential to produce Type I errors. There were 16 different outcome measures (4 primary and 12 secondary) in the study and multiple statistical tests associated with this large number of outcomes. However, it is argued that the complexity of the Mate-Tricks program utilizing both social cognitive and social ecological frameworks necessitated a wide spread of measures looking at a range of outcomes from multiple perspectives in order to get a clear picture of the program's effects. One solution to reducing the likelihood of Type I error is to adjust the significance level for the number of tests; others, however, argue that these practices in educational studies may not be necessary and may lead to an increase in Type II errors (Gelman, Hill, & Yajima, 2012). It is also argued that an adjustment for clustering is more important for reducing statistical error in education studies, which has been done in this study by reporting robust standard errors. A second limitation is missing data that may lead to a reduction in study power, which could potentially lead to Type II errors. However, imputation of pretest data allowed all posttest data to be included in the analysis, which substantially improved study power. So, overall, it is argued that the number and diversity of study outcomes under investigation in this study provided a holistic view of the complex program effects, and this was combined with consistent patterns of significant negative effects on behavioral and parenting outcomes. This would suggest that Type I and Type II errors were not a major problem in this study. The final limitation is the consistency of study effects between child and teacher reports; that is, all the significant effects were child reported. Possible reasons for this would require further research, but there may have been an issue with the variation of teacher responses and as a result their sensitivity to change. A small number of teachers filled out questionnaires for a comparatively large number of children (all the children in their respective classes), whereas each

child responded from an individual perspective. This notion is supported by the fact that teacher-reported standard deviations and robust standard errors for the outcome measures are frequently close to or higher than their respective mean scores and coefficients, which would indicate that teacher measures are less representative of outcomes than the child-reported measures.

In conclusion, the major theoretical contribution of this article is that the demographic characteristics of the children (e.g., gender, affluence, and special educational needs) were not found to be greatly influential on the effects of this after-school prosocial behavior program. However, implementation factors were found to have significant associations with program effects. Cohort effects, and in particular parental participation, were found to be significant predictors of the effectiveness of the program. Furthermore, there is a clear gap in the research regarding the association between parenting components of SEL programs and program outcomes both in school and after school. Basically, this study provides evidence that the relationship is significant but complex, and substantial further research is required. Overall, the study showed that the program produced several negative effects even when program quality was high and there were strong partnerships between schools, communities, and with many families. In fact, the evidence from this research, and that drawn from other rigorous studies (James-Burdumy et al., 2008; Linden et al., 2011), would suggest that group-based after-school social learning programs have real potential to cause iatrogenic effects and therefore must be designed, piloted, evaluated, and implemented with a high degree of care.

## Notes

Liam O'Hare, Andy Biggart, Karen Kerr, and Paul Connolly are with the Centre for Effective Education, School of Education, Queen's University Belfast. The research was funded by Atlantic Philanthropies. Correspondence should be directed to Liam O'Hare; e-mail: l.ohare@qub.ac.uk

1. These were U.S. RCT studies of nontargeted after-school programs reporting behavioral, social, and emotional or academic outcomes for children and young people aged 5–19.
2. Supportive parenting is deemed positive, while authoritarian and liberal parenting are undesirable parenting styles.
3. At the request of the implementation team, there was only one testing session for each school in order not to overburden schools with testing.
4. For transparency, none of the parental measures showed significant effects of the intervention on any outcomes. However, the small sample size had low analytic power.

## References

- Achenbach, T. M., Becker, A., Döpfner, M., Heiervang, E., Roessner, V., Steinhausen, H. C., & Rothenberger, A. (2008). Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: Research findings, applications, and future directions. *Journal of Child Psychology and Psychiatry*, *49*(3), 251–275.
- Adelman H. S., & Taylor, L. (2003). On sustainability of project innovations as systemic change. *Journal of Educational and Psychological Consultation*, *14*(1), 1–25.
- American Educational Research Association. (2011). *Code of Ethics*. American Educational Research Association, approved by the AERA Council February 2011. Retrieved from [http://www.aera.net/Portals/38/docs/About\\_AERA/CodeOfEthics%281%29.pdf](http://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics%281%29.pdf)
- Biggart, A., Kerr, K., O'Hare, L., & Connolly, P. (2013). A randomised control trial evaluation of a literacy after-school programme for struggling beginning readers. *International Journal of Educational Research*, *62*, 129–140.



- Bool, M., & Onrust, S. (2009). Gezin aan Bod: Het Strengthening Families Program. *Verslaving*, 5(4), 21–32.
- Botvin, G. J., & Griffin, K. W. (2002). Life skills training as a primary prevention approach for adolescent drug abuse and other problem behaviors. *International Journal of Emergency Mental Health*, 4, 41–47.
- British Educational Research Association. (2004). *Revised ethical guidelines for educational research*. Southwell, Notts: BERA. Retrieved from <http://www.bera.ac.uk/system/files/ethica1.pdf>
- Bronfenbrenner, U., & Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Center for the Study of Violence Prevention. (2011). *Blueprints*. University of Colorado. Retrieved from <http://www.colorado.edu/cspv/blueprints/>
- Childhood Development Initiative (CDI). (2010). *Mate-Tricks programme manual*. Dublin: Tal-laght West Childhood Development Initiative.
- Cross, M., Gottfredson, D., Wilson, D., Rorie, M., & Connell, N. (2010). Implementation quality and positive experiences in after school programs. *American Journal of Community Psychology*, 45, 370–380.
- Davis, F., McDonald, L., & Axford, N. (2012). *Technique is not enough: A framework for ensuring that evidence-based parenting programmes are socially inclusive*. London: British Psychological Society. Retrieved from <http://www.bps.org.uk/system/files/images/tine.pdf>
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Ser. B (Methodological)*, 1–38.
- Derzon, J., Sale, E., Springer, J., & Brounstein, P. (2005). Estimating intervention effectiveness: Synthetic projection of field evaluation results. *Journal of Primary Prevention*, 26, 321–343.
- Diekstra, R. F. (2008). Effectiveness of school-based social and emotional education programmes worldwide. *Social and emotional education: An international analysis*, 255–312.
- Dishion, T. J., & Kavanagh, K. (2003). *Intervening in adolescent problem behavior: A family-centered approach*. New York: Guilford.
- Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., . . . Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology*, 14(3), 317–345.
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., . . . Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1(3), 6–28.
- Dumas, J., Lynch, A., Laughlin, J., Smith, E., & Prinz, R. (2001). Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the Early Alliance prevention trial. *American Journal of Preventive Medicine*, 20(1), 38–47.
- DuPaul, G. J. (1991). Parent and teacher ratings of ADHD symptoms: Psychometric properties in a community based sample. *Journal of Clinical Child Psychology*, 20, 245–253.
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 45, 294–309.
- Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research*, 20(3), 308–313.
- Dusenbury, L., & Hansen, W. B. (2004). Pursuing the course from research to practice. *Prevention Science*, 5(1), 55–59.

- Erikson, J. (2002). *Effects of the Strengthening Families Program with Native American families*. Paper presented at the International Nursing Research Congress, Sigma Theta Tau International, July 2002.
- Farrell, A., Meyer, A., Kung, E., & Sullivan, T. (2001). Development and evaluation of school-based violence prevention programs. *Journal of Clinical Child Psychology*, *30*(1), 207–220.
- Fossum, S., Handegard, B., Martinussen, M., & Morch, W. (2008). Psychosocial interventions for disruptive and aggressive behavior in children and adolescents: A meta-analysis. *European child and adolescent psychiatry*, *17*(7), 438–451.
- Frick, P. J. (1991). *The Alabama Parenting Questionnaire*. Unpublished rating scale, University of Alabama.
- Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, *5*(2), 189–211.
- Gottfredson, D., Cross, A. B., Wilson, D., Rorie, M., & Connell, N. (2010). Effects of participation in after-school programs for middle school students: A randomized trial. *Journal of Research on Educational Effectiveness*, *3*(3), 282–313.
- Gottfredson, D., & Gottfredson, G. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, *39*(1), 3–35.
- Gottfredson, D., Kumpfer, K., Polizzi-Fox, D., Wilson, D., Puryear, V., Beatty, P., & Vilmenay, M. (2006). The Strengthening Washington D.C. Families Project: A randomized effectiveness trial of family-based prevention. *Prevention Science*, *7*(1), 57–74.
- Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and Psychopathology*, *7*, 117–136.
- Hallfors, D., & Godette, D. (2002). Will the “principles of effectiveness” improve prevention practice? Early findings from a diffusion study. *Health Education Research*, *17*, 461–470.
- Hutchings, J. (2012). From ABA to SPR: 30 years developing evidence based services for the treatment and prevention of conduct disorder in Wales. *Journal of Children's Services*, *7*(2), 101–112.
- James-Burdumy, S., Dynarski, M., & Deke, J. (2008). After-school program effects on behavior: Results from the 21st Century Community Learning Centers Program national evaluation. *Economic Inquiry*, *46*(1), 13–18.
- January, A., Casey, R., & Paulson, D. (2011). A Meta-analysis of classroom-wide interventions to build social skills: Do they work? *School Psychology Review*, *40*(2), 242–256.
- Kehoe, S., & O'Hare, L. (2010). The reliability and validity of the Family Affluence Scale. *Effective Education*, *2*(2), 155–164.
- Kumpfer, K. L., & Alvarado, R. (2003). Family strengthening approaches for the prevention of youth problem behaviors. *American Psychologist*, *58*, 457–465.
- Kumpfer, K. L., Greene, J., Whiteside, H., & Allen, K. (2010). Effectiveness outcomes of four age versions of the Strengthening Families Programme in state-wide field sites. *Group Dynamics: Theory, Research and Practice*, *14*(3), 211–229.
- Kumpfer, K. L., & Turner, C. W. (1991). The social ecology model of adolescent substance abuse: Implications for prevention. *International Journal of the Addictions*, *25*, 435–463.
- Lauver, S. (2002). *Assessing the benefits of an after-school program for urban youth: Results of an experimental design and process evaluation*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.
- Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, *2*(1), 53–70.
- Linden, L. L., Herrera, C., & Grossman, J. B. (2011). *Achieving academic success after school: A randomized evaluation of the Higher Achievement Program*. Retrieved from [http://www.wallacefoundation.org/knowledge-center/summer-and-extended-learning-time/summer-learning/Documents/Achieving-Academic-Success-After-School-Technical\\_Report.pdf](http://www.wallacefoundation.org/knowledge-center/summer-and-extended-learning-time/summer-learning/Documents/Achieving-Academic-Success-After-School-Technical_Report.pdf)
- Lipsey, M. W. (1992). The effect of treatment on juvenile delinquents: Results from meta-analysis. *Psychology and Law: International Perspectives*, 131–143.
- Little, P., Wimer, C., & Weiss, H. B. (2008). After school programs in the 21st century: Their potential and what it takes to achieve it. *Issues and Opportunities in Out-of-School Time Evaluation*, *10*, 1–12.

- Lochman, J., Boxmeyer, C., Powell, N., Wells, K., & Windle, M. (2009). Dissemination of the Coping Program: Importance of intensity of counsellor training. *Journal of Counselling and Clinical Psychology*, 77(3), 397–409.
- Lochman, J. E., & Wells, K. C. (2002a). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the Coping Power Program. *Development and Psychopathology*, 14(4), 945–967.
- Lochman, J. E., & Wells, K. C. (2002b). The Coping Power Program at the middle-school transition: Universal and indicated prevention effects. *Psychology of Addictive Behaviors*, 16(4S), 40–54.
- Lochman, J. E., & Wells, K. C. (2004). The Coping Power Program for preadolescent aggressive boys and their parents: Outcome effects at the 1-year follow-up. *Journal of Consulting and Clinical Psychology*, 72(4), 571.
- Lösel, F., & Beelmann, A. (2003). Effects of child skills training in preventing antisocial behavior: A systematic review of randomised evaluations. *Annals of the American Academy of Political and Social Science*, 587(1), 84–109.
- Maas, C. J., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, 58(2), 127–137.
- Mavroveli, S., Petrides, K. V., Rieffe, C., & Bakker, F. (2007). Trait emotional intelligence, psychological well-being and peer-related social competence in adolescence. *British Journal of Developmental Psychology*, 25, 263–275.
- McCart, M., Priester, P., Davis, W., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34(4), 527–543.
- O'Hare, L., Kerr, K., Biggart, A., & Connolly, P. (2012). *Evaluation of the effectiveness of the Childhood Development Initiative's Mate-Tricks Pro-Social Behaviour After-School Programme*. Dublin: Childhood Development Initiative (CDI).
- Olweus, D., Limber, S., & Mihalic, S. F. (1999). *Blueprints for violence prevention, book nine: Bullying prevention program*. Boulder, CO: Center for the Study and Prevention of Violence.
- Orte, C., Touza, C., Ballester, L., & March, M. (2008). Children of drug-dependent parents: Prevention programme outcomes. *Educational Research*, 50(3), 249–260.
- Piehler, T. F., & Dishion, T. J. (2007). Interpersonal dynamics within adolescent friendships: Dyadic mutuality, deviant talk, and patterns of antisocial behavior. *Child Development*, 78(5), 1611–1624.
- Quinn, M., Kavale, K., Mathur, S., Rutherford, R., & Forness, S. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders. *Journal of Emotional and Behavioural Disorders*, 7(1), 54–64.
- Reddy, L. A., Newman, E., De Thomas, C. A., & Chun, V. (2009). Effectiveness of school-based prevention and intervention programs for children and adolescents with emotional disturbance: A meta-analysis. *Journal of School Psychology*, 47(2), 77–99.
- Riesch, S. K., Brown, R. L., Anderson, L. S., Wang, K., Cauty-Mitchell, J., & Johnson, D. L. (2012). Strengthening Families Program (10–14): Effects on the family environment. *Western Journal of Nursing Research*, 34(3), 340–376.
- Rigby, K., & Slee, P. T. (1992). Dimensions of interpersonal relations among Australian school children and implications for psychological well-being. *Journal of Social Psychology*, 133, 33–42.
- Rothbaum, F., & Weisz, J. R. (1994). Parental caregiving and child externalizing behavior in non-clinical samples: A meta-analysis. *Psychological bulletin*, 116(1), 55.
- Schnohr, C. W., Kreiner, S., Due, E. P., Currie, C., Boyce, W., & Diderichsen, F. (2008). Differential item functioning of a family affluence scale: Validation study on data from HBSC 2001/02. *Social Indicators Research*, 89(1), 79–95.
- Semeniuk, Y., Brown, R. L., Riesch, S. K., Zywicki, M., Hopper, J., & Henriques, J. B. (2010). The Strengthening Families Program 10–14: Influence on parent and youth problem-solving skill. *Journal of Psychiatric and Mental Health Nursing*, 17(5), 392–402.
- Skärstrand, E., Sundell, K., & Andréasson, S. (2010). *Evaluation of a Swedish version of the Strengthening Families Programme: Three-year outcomes of a cluster randomized trial. Prevention of alcohol and drug problems among adolescents: Evaluating a Swedish version of the Strengthening*

- Families Program*. Stockholm: Karolinska Institutet, Department of Public Health Sciences, Division of Social Medicine.
- Thaker, S., Steckler, A., Sanchez, V., Khatapoush, S., Rose, J., & Hallfors, D. (2008). Programme characteristics and organisational factors affecting the implementation of a school-based indicated prevention program. *Health Education Research*, *23*(2), 238–248.
- U.S. Department of Education, Office of the Under Secretary. (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers Program, first year findings*. Washington, DC: U.S. Department of Education.
- Webster-Stratton, C. (2012). *The Incredible Years Universal Parenting Program—Attentive Parenting*. Retrieved from <http://www.incredibleyears.com/program/attentive-parenting-program-delivery.pdf>
- Webster-Stratton, C., & Reid, M. J. (2003). *The Incredible Years parent, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems*. Maryland After School Community Grant Program: Report on the 1999–2000 school year evaluation of the phase 1 after-school programs, University of Maryland, College Park.
- Wilson, S., Lipsey, M., & Derzon, J. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis, *Journal of Consulting and Clinical Psychology*, *71*(1), 136–149.
- Zief, S. G., Lauver, S., & Maynard, R. A. (2006). Impacts of after-school programs on student outcomes. *Campbell Systematic Reviews*, *3*.